American Journal of Epidemiology
Vol. 152, No. 3
Printed in U.S.A.


# Evaluation of Alternative Approaches to Assign Nutrient Values to Food Groups in Food Frequency Questionnaires

Amy F. Subar,[1] Douglas Midthune,[1] Martin Kulldorff,[2] Charles C. Brown,[1] Frances E. Thompson,[1] Victor Kipnis,[1] and Arthur Schatzkin[1]

Although every food frequency questionnaire (FFQ) requires a nutrient database to produce nutrient intake estimates, it is often unclear how a particular database has been generated. Moreover, alternative methods for constructing a database have not been rigorously evaluated. Using 24-hour recalls from the 1994–1996 Continuing Survey of Food Intake by Individuals, the authors categorized 5,261 individual foods reported by 10,019 adults into 170 food groups consistent with line items on an FFQ. These food groups were used to generate 10 potential nutrient databases for a FFQ that varied by whether the authors 1) used means or medians, 2) did or did not consider age, 3) incorporated collapsing strategies for small age-gender-portion size cells, 4) excluded outliers in a regression, and 5) used weighted median nutrient density × age-gender-portion size-specific median gram weights (Block method). Mean error, mean squared error, and mean absolute error were calculated and compared across methods, with error being the difference in total observed (from recalls for each individual) and total estimated intake (from each of the 10 methods) for seven nutrients. Mean methods for assigning nutrients to food groups were superior to median approaches for all measurements. Among the mean methods, no single variation was consistently better. *Am J Epidemiol* 2000;152:279–86.

diet; epidemiologic methods; food habits; nutrition assessment; nutrition surveys; questionnaires

Food frequency questionnaires (FFQs) are often used to measure usual dietary intake in nutritional epidemiologic research. FFQs require algorithms that convert reported frequency of consumption and, in some cases, portion size for each food item into nutrient values that can then be summed across foods to estimate average total daily nutrient intake. Compared with other self-reported dietary assessment methods, such as 24-hour dietary recalls or food records, FFQs obtain less detailed information on food type or portion size. Therefore, the nutrients assigned to each line item on an FFQ necessarily represent composite values for a number of possible variants of the food item queried. For example, the FFQ food item, "French fries, hash browned, or other fried potatoes" is a composite of several types of fried potatoes with varying levels of nutrient content. Thus, the method by which nutrients are assigned to FFQ line items require careful thought and informed decision making.

The Block and Willett FFQs, or variations of either, are two examples of FFQs that are widely used in nutritional epidemiology research. Block et al. (1) pioneered a data-driven approach, which used food consumption data to develop an FFQ and its associated analytical software. Dietary intake data from a large, nationally representative sample (1976–1980 National Health and Nutrition Survey II) were used to decide which foods to include in the FFQ and to assign nutrient composition and portion sizes. To create a food list for an FFQ, Willett et al. (2) used regression methods (using food record data) and judgment to prepare an extensive list of commonly consumed foods containing nutrients pertinent to the prevention of cancer and heart disease. In the Willett FFQ, portion size is not specifically asked of respondents, but within the frequency question, respondents are asked how often a particular standard portion size is consumed. Judgment is used to establish both the size of these standard portion sizes (in common household units) and the nutrient content of the FFQ line items (using current food composition databases) (3). Various other methods of assigning nutrient composition values to line items on an FFQ are documented in the literature (4–6). However, there has been no evaluation of the relative performance of the methods.

Several scientific bodies have made strong recommendations for improving dietary assessment methods (7, 8). In this paper, we investigate whether one aspect of the overall validity of an FFQ, the assignment of nutrient values, is sensitive to different assignment methods. The research was motivated by a need to provide a nutrient database for new FFQs developed and used in research at the National Cancer Institute (NCI). Specifically, we calculated nutrient values for food items queried on the FFQ used in the National

Received for publication June 1, 1999, and accepted for publication October 15, 1999.

Abbreviations: AARP-FFQ, American Association of Retired Persons food frequency questionnaire; CSFII, Continuing Survey of Food Intakes by Individuals; FFQ, food frequency questionnaire; NCI, National Cancer Institute.

[1] National Cancer Institute, Bethesda, MD.
[2] University of Connecticut Medical School, Farmington, CT.
Reprint requests to Dr. Amy F. Subar, National Cancer Institute, 6130 Executive Blvd., MSC 7344, EPN 4005, Bethesda, MD 20892–7344 (e-mail: amy_subar@nih.gov).

Institutes of Health American Association of Retired Persons Diet and Health Study. However, the results are relevant to FFQs in general.

## MATERIALS AND METHODS

### Continuing Survey of Food Intake by Individuals (CSFII)

In developing the American Association of Retired Persons food frequency questionnaire (AARP-FFQ) and its associated nutrient database, we used data from the United States Department of Agriculture's 1994–1996 CSFII. In each of three CSFII survey years, a nationally representative sample of noninstitutionalized individuals who resided in all 50 states provided, through in-person interviews, a 24-hour dietary recall on two nonconsecutive days 3–10 days apart. The survey included an oversampling of the low-income population. Data were collected from a selected individual within each household. The 24-hour dietary recall included multiple passes through the list of all foods and beverages recalled by the respondent in order to maximize the amount of information collected. Further information regarding the sampling procedures and the data collection is provided elsewhere (9). All statistical analyses used weighting factors designed to provide nationally representative estimates and to adjust for differential rates of selection and nonresponse.

Eighty percent of the eligible respondents participated in the first interview, and 95 percent of those who responded to the first interview participated in the second interview, resulting in an overall response rate of 76 percent for both days of data collection. This analysis includes 10,019 adult respondents aged 19 years or older who completed either 1 or 2 days of dietary recall.

### FFQ line items

We categorized the 5,261 individual food codes found in the CSFII database and consumed by these adults into 182 food groups similar in usage and nutrient content. For example, a ready-to-eat cereal group was created from the 111 individual food codes for all types of ready-to-eat cold breakfast cereals reported. Analyses of important food sources of nutrients, similar to those reported by others (10–14), were conducted to assess which of the 182 food groups were important food sources of energy, fat, percentage of energy from fat, vitamin C, beta-carotene, dietary fiber, vitamin A, calcium, and vitamin E. The 124 groups that contributed 90 percent or more of the total intake for each of these nutrients/food constituents were selected to create the final food list for the AARP-FFQ. The other 58 food groups were excluded because they contributed little to nutrient intake in the United States, usually because of infrequent consumption. For some food groups, subgroups were created to better assess varying nutrient contents within the broader food group, especially for fat and fiber. For example, the ready-to-eat cereals food group consisted of four subgroups to differentiate the cereals more clearly with respect to fiber and other nutrients (highly fortified, very high fiber, moderate fiber, and other). The FFQ first queries frequency of total cereal intake and usual portion size and then presents the embedded questions regarding the proportion of the time each of the four different varieties is consumed. Our decision to create subgroups for a large food group, as opposed to creating separate food groups, was based on cognitive research indicating that respondents had difficulty when asked to complete the frequency of each of many related items (15). Including subgroups, nutrient content estimation was needed for 170 food groups.

### FFQ portion size

The AARP-FFQ developed at the NCI retains questions about portion size. In specifying the portion size options for FFQ line items, we adopted line item-specific ranges, such as "less than 1 cup," "1–2 cups," and "greater than 2 cups," rather than "small," "medium," and "large." We did this because results from cognitive testing for the new FFQ (15) suggested that study participants were more able to answer questions about portion size when quantified range options were provided.

To establish portion size ranges using CSFII, we first looked at nationally representative data from a 68-item NCI Block FFQ administered in the 1992 National Health Interview Survey (16). The data showed that, across all food items, portion size was answered as "medium" about 66 percent of the time. We therefore decided that the range for the middle portion size should be broader than the middle third of the CSFII portion size distribution to better represent the tendency for individuals to select a middle portion size on FFQs. We selected the approximate 25th and 75th percentiles of gram weight portion sizes for each food group as cutpoints to define our three portion size ranges. This created a broad medium portion size, but left enough CSFII respondents in the small and large portion size groups to provide stable estimates of the amounts consumed.

Because portion size and types of foods consumed varies by age and gender, we separated respondents into three age groups (19–30, 31–50, and >50 years) by gender to assess age-specific portion sizes. We chose these age groups because they are similar to those upon which the Recommended Dietary Allowance energy requirements are based (19–24, 25–50, and >50 years) (17), yet allow for adequate numbers of CSFII respondents for analyses in the youngest age group.

### Deriving nutrient estimates: alternative approaches

In general, the average daily intake of a given nutrient for an FFQ respondent is derived as follows: For each line item, the reported daily frequency is multiplied by a nutrient value specific to the respondents age, gender, and reported portion size. The calculated line item-specific values are then summed across all the line items, yielding a total nutrient intake for that respondent.

The source of methodological variability we sought to investigate was the derivation of the line item-specific nutrient values. Because the general method of calculating nutrient intake (outlined above) requires a nutrient value (age-, gender-, and portion size-specific) for each line item, the challenge is to assign a single nutrient value, given that respondents eat a variety of different foods subsumed by that line item. For example, people eat a variety of cheeses, each with its own dietary fat content. What fat value should one assign for the line item "cheese"? We used the CSFII, which provides detailed, nationally representative food consumption data regarding the intake of many varieties of food (cheese types, for example) to evaluate alternative methods of assigning nutrient values for a single line item. In all analyses, we used 24-hour recall data from CSFII; no FFQ nutrient data were analyzed.

We first categorized food consumptions of all the CSFII respondents as reported on the 24-hour recalls into $3 \times 2 \times 3 = 18$ age-, gender-, and portion size-specific cells for each of the 170 food groups being evaluated. We computed means and medians for each of these 18 categories for each of seven nutrients or dietary constituents representing a variety of macro- and micronutrients (energy, fat, carbohydrate, fiber, vitamin A, vitamin C, and iron). We investigated a series of methodological variations. First, to investigate the influence of small numbers, we examined the effect of collapsing the data over adjacent age groups within gender and portion size when there were fewer than 10 individuals in one of the 18 cells. Second, to investigate the importance of age in determining nutrient intake, we combined the three age groups by gender and portion size before computing means and medians, thereby reducing the number of cell sizes from 18 to six. Third, we examined a simple gender-specific regression approach—nutrient = age effect + portion size effect—and then repeated the regression excluding outliers to address the problem of dietary data being skewed by high intakes. Fourth, we computed estimates by using the method developed by Block et al. (1), in which a single median nutrient density (weighted by frequency of consumption of individual foods within a food group) is multiplied by a median age- and gender-specific gram weight portion size.

We thus compared 10 nutrient estimation methods, by gender, characterized as follows: 1) mean nutrient by portion size and age; 2) mean nutrient by portion size and age, with collapsing; 3) mean regression of nutrient on portion size and age; 4) mean nutrient by portion size; 5) mean regression of nutrient on portion size and age, excluding outliers; 6) median nutrient by portion size and age; 7) median nutrient by portion size and age, with collapsing; 8) median regression of nutrient on portion size and age; 9) median nutrient by portion size; 10) median nutrient density (weighted by frequency of consumption of individual foods within a food group) × median age- and gender-specific portion size (Block method).

For all regressions, the model was: nutrient intake $= \beta_0 + \beta_1\,\text{Age2} + \beta_2\,\text{Age3} + \beta_3\,\text{Size2} + \beta_4\,\text{Size3} + \varepsilon$, where $\text{Age}_i$ is an indicator for age group $i$ and $\text{Size}_i$ is an indicator for portion-size group $i$. A regression model was fit separately for each gender group. In model 5, outliers from the model 3 regression were defined as observations with squared errors greater than three times the mean squared error.

We then sought to compare the 10 methods described above to determine which performed best. For each of 170 food groups corresponding to FFQ line items, we now had 10 different nutrient values from our estimation methods based on the grouped CSFII data analyses (described above) for each of the seven nutrients. We compared these nutrient values with those for each specific food item reported on 24-hour recalls (among our 170 food groups) for all CSFII respondents. Specifically, we calculated the difference between each nutrient value for each food reported on an individual's 24-hour recall (observed value) and the nutrient value for each of our 10 methods (estimated values). We then created an error term by summing these differences by nutrient for all foods reported in each individual's 24-hour dietary recall to evaluate error in terms of total daily nutrient intake.

We evaluated our 10 methods on the basis of three measures of error estimation: mean error, mean squared error, and mean absolute error. Mean error measures the magnitude and direction of the possible bias of the estimate, while mean squared error and mean absolute error measure the precision of the estimate (bias + variation). Estimators that minimize mean absolute error are sometimes preferred over those that minimize mean squared error because the latter can be sensitive to outliers.

To compare methods, we examined the associated errors of each one. Generally, when the same data are used both to fit a model and to estimate errors, the usual estimated errors, or residuals, tend to be biased toward zero. This bias depends on the complexity of the model and the method of fitting it. Therefore, we used cross-validation to obtain unbiased error estimates that could be compared across methods. This applied cross-validation method leaves one subject out of the data, predicts that subject's nutrient intake based on the remaining $n - 1$ subjects, and calculates error for that subject by subtracting the actual value from the predicted value. This was done for each subject one at a time.

## RESULTS

Table 1 presents, for seven nutrients or dietary constituents, the mean estimated total nutrient values for men aged 31–50 years only (data for all other age-gender groups available upon request). The magnitude of differences in nutrient values estimated between the 10 different methods was generally between 10 and 15 percent. Across all nutrients, mean methods tended to produce higher estimates than median methods. Among the mean methods, that which excluded outliers tended to produce the lowest estimates. The Block method estimates tended to be slightly higher than those from the median method for most nutrients but lower than those from the mean methods. Results were similar for respondents aged 19–30 and those aged more than 50 years (data not shown).

Mean error, shown in table 2, is a measure of the average bias of the 10 estimates. With the exception of the regression method excluding outliers, the mean methods were consistently better than the median or Block approaches, showing little bias. The regression method that excludes outliers exhibited the greatest bias among the mean methods. The median and Block methods all had comparable biases.

**TABLE 1. Estimated mean and median nutrient intakes from 10 estimation methods by gender, adults aged 31–50 years, CSFII*, 1994–1996**

| Methods | Energy (kcal) | Fat (g) | Carbohydrate (g) | Fiber (g) | Vitamin A (RE*) | Vitamin C (mg) | Iron (mg) |
|---|---|---|---|---|---|---|---|
| *Men* | | | | | | | |
| Mean | | | | | | | |
| By gender | 1,930 | 70.0 | 236.2 | 14.0 | 800.6 | 77.4 | 14.2 |
| By gender and portion size | 1,938 | 69.8 | 238.6 | 13.9 | 823.4 | 78.2 | 14.2 |
| By age, gender, and portion size | 1,945 | 70.3 | 239.5 | 14.1 | 820.7 | 78.2 | 14.4 |
| By age, gender, and portion size/collapsing† | 1,946 | 70.4 | 239.3 | 14.1 | 826.2 | 78.1 | 14.4 |
| Regression‡ | 1,945 | 70.3 | 239.4 | 14.1 | 821.9 | 78.1 | 14.4 |
| Regression excluding outliers§ | 1,833 | 65.8 | 225.2 | 13.1 | 750.3 | 72.3 | 13.4 |
| Median | | | | | | | |
| By gender and portion size | 1,762 | 63.1 | 217.2 | 12.5 | 702.1 | 64.6 | 12.6 |
| By age, gender, and portion size | 1,767 | 63.9 | 216.3 | 12.6 | 710.8 | 65.6 | 12.8 |
| By age, gender, and portion size/collapsing† | 1,766 | 63.8 | 216.3 | 12.6 | 708.4 | 65.3 | 12.7 |
| Regression‡ | 1,764 | 63.2 | 218.2 | 12.6 | 696.5 | 65.6 | 12.7 |
| Block¶ | 1,778 | 63.8 | 215.2 | 13.0 | 701.7 | 66.7 | 13.2 |
| *Women* | | | | | | | |
| Mean | | | | | | | |
| By gender | 1,255 | 44.1 | 163.7 | 10.0 | 655.6 | 66.6 | 9.6 |
| By gender and portion size | 1,258 | 43.9 | 164.9 | 10.0 | 660.4 | 66.6 | 9.7 |
| By age, gender, and portion size | 1,258 | 44.1 | 164.4 | 10.1 | 657.7 | 65.9 | 9.7 |
| By age, gender, and portion size/collapsing† | 1,260 | 44.2 | 164.6 | 10.1 | 662.0 | 66.0 | 9.7 |
| Regression‡ | 1,258 | 44.1 | 164.3 | 10.1 | 657.8 | 66.0 | 9.7 |
| Regression excluding outliers§ | 1,214 | 42.2 | 157.9 | 9.6 | 615.0 | 62.1 | 9.2 |
| Median | | | | | | | |
| By gender and portion size | 1,184 | 40.3 | 154.7 | 9.2 | 564.9 | 57.6 | 8.7 |
| By age, gender, and portion size | 1,189 | 40.6 | 154.9 | 9.2 | 562.8 | 57.6 | 8.8 |
| By age, gender, and portion size/collapsing† | 1,190 | 40.7 | 155.0 | 9.2 | 566.2 | 57.8 | 8.8 |
| Regression‡ | 1,184 | 40.3 | 154.6 | 9.2 | 563.2 | 57.5 | 8.8 |
| Block¶ | 1,196 | 41.2 | 153.9 | 9.2 | 574.8 | 57.2 | 9.3 |

* CSFII, Continuing Survey of Food Intakes by Individuals; RE, retinol equivalents.

† If there were fewer than 10 individuals within any age, gender, and portion size strata, adjacent age group strata were collapsed.

‡ Regression model: nutrient intake = $\beta_0 + \beta_1$ Age2 + $\beta_2$ Age3 + $\beta_3$ Size2 + $\beta_4$ Size3 + $\varepsilon$, where $\text{Age}_i$ is an indicator for age group *i* and $\text{Size}_i$ is an indicator for portion-size group *i*.

§ Outliers are defined as observations having squared errors greater than three times the mean squared error.

¶ Weighted median nutrient density × age-gender-portion size-specific median gram weights.

Table 3 shows the mean absolute error, a measure of the precision of the estimate, of each of the 10 methods. For energy, fat, fiber, and iron, mean methods produced the smallest error, followed by median methods, followed by the Block method, although the differences were particularly small for iron. Results for mean and median methods for carbohydrate, vitamin A, and vitamin C were equivocal. The Block method produced slightly more error than did the other methods. Within either the mean or the median approach, mean absolute errors between methods were within 1 percent of each other for all nutrients. The regression approach, excluding outliers, was the most consistently better mean approach, but by very little. No approach was consistently better for the median methods.

Results for mean squared error are presented in table 4. For all nutrients except energy for men, the Block method produced the largest mean squared error across all methods. The mean methods had a smaller mean squared error than did the median methods, and within mean or median methods, mean squared error values were similar, with no single approach being consistently better.

## DISCUSSION

This research begins with the premise, pioneered by Block et al. (1), that using nationally representative dietary data in a systematic way to create an FFQ nutrient

**TABLE 2. Mean error between reported nutrient intakes from 24-hour recalls and 10 estimation methods by gender, adults aged 19 years or more, CSFII*, 1994–1996**

| Methods | Energy (kcal) | Fat (g) | Carbo-hydrate (g) | Fiber (g) | Vitamin A (RE*) | Vitamin C (mg) | Iron (mg) |
|---|---|---|---|---|---|---|---|
| *Men* | | | | | | | |
| Mean | | | | | | | |
| By gender and portion size | –24 | –0.49 | –4.1 | –0.15 | –15 | –0.9 | –0.16 |
| By age, gender, and portion size | –24 | –0.51 | –4.0 | –0.15 | –12 | –0.7 | –0.14 |
| By age, gender, and portion size/collapsing† | –24 | –0.49 | –4.0 | –0.15 | –13 | –0.8 | –0.15 |
| Regression‡ | –24 | –0.50 | –3.9 | –0.14 | –12 | –0.7 | –0.14 |
| Regression excluding outliers§ | 73 | 3.37 | 8.2 | 0.63 | 56 | 5.6 | 0.66 |
| Median | | | | | | | |
| By gender and portion size | 144 | 5.89 | 16.8 | 1.23 | 118 | 13.7 | 1.54 |
| By age, gender, and portion size | 138 | 5.78 | 16.5 | 1.18 | 109 | 12.1 | 1.41 |
| By age, gender, and portion size/collapsing† | 139 | 5.90 | 16.6 | 1.19 | 109 | 12.2 | 1.43 |
| Regression‡ | 144 | 6.03 | 16.6 | 1.19 | 121 | 13.2 | 1.47 |
| Block¶ | 136 | 6.09 | 19.6 | 0.99 | 131 | 12.4 | 0.92 |
| *Women* | | | | | | | |
| Mean | | | | | | | |
| By gender and portion size | –7 | –0.15 | –1.3 | –0.02 | –8 | –0.2 | –0.05 |
| By age, gender, and portion size | –7 | –0.15 | –1.3 | –0.02 | –9 | –0.1 | –0.04 |
| By age, gender, and portion size/collapsing† | –7 | –0.16 | –1.3 | –0.02 | –8 | –0.1 | –0.04 |
| Regression‡ | –7 | –0.14 | –1.3 | –0.02 | –8 | –0.1 | –0.03 |
| Regression excluding outliers§ | 36 | 1.60 | 4.9 | 0.38 | 37 | 4.0 | 0.37 |
| Median | | | | | | | |
| By gender and portion size | 62 | 3.24 | 8.3 | 0.75 | 88 | 9.3 | 1.00 |
| By age, gender, and portion size | 61 | 3.25 | 8.0 | 0.76 | 85 | 9.4 | 0.89 |
| By age, gender, and portion size/collapsing† | 62 | 3.27 | 8.1 | 0.77 | 89 | 9.5 | 0.91 |
| Regression‡ | 63 | 3.22 | 8.4 | 0.77 | 89 | 9.4 | 0.94 |
| Block¶ | 62 | 2.95 | 10.3 | 0.87 | 91 | 11.3 | 0.33 |

* CSFII, Continuing Survey of Food Intakes by Individuals; RE, retinol equivalents.

† If there were fewer than 10 individuals within any age, gender, and portion size strata, adjacent age group strata were collapsed.

‡ Regression model: nutrient intake = $\beta_0 + \beta_1$ Age2 + $\beta_2$ Age3 + $\beta_3$ Size2 + $\beta_4$ Size3 + $\varepsilon$, where $Age_i$ is an indicator for age group *i* and $Size_i$ is an indicator for portion-size group *i*.

§ Outliers are defined as observations having squared errors greater than three times the mean squared error.

¶ Weighted median nutrient density × age-gender-portion size-specific median gram weights.

database will provide the most unbiased nutrient estimates for an FFQ to be used among US adults. This reasoning is based on the premise that non-data-driven methods will necessarily involve more extensive judgments and assumptions on the part of investigators that may not accurately represent current food consumption and composition and that are not reproducible across investigators. Although all methods require some degree of judgment and decision making (such as in the food grouping step), the more this is minimized, the more likely that the instruments will reflect the reality of consumption in the population.

The findings show that using either the mean or the median nutrient intakes of all reports within a given portion size for a given food group is an improvement over the current Block approach. Because there is little or no documentation on how nutrient databases are constructed for other FFQs, we are unable to evaluate how our newer methods might compare with them. Further, the findings indicate clearly that data-driven methods that use mean versus median approaches are superior with respect to mean bias, absolute error, and mean squared error of total daily nutrient intake.

An unexpected finding was that age was not a critical variable in nutrient estimation. Although nutrient intake and food choices are known to vary by age and gender, these data show that, once portion sizes are defined by the 25th and 75th percentiles of gram weight intakes for all

**TABLE 3. Mean absolute error between reported nutrient intakes from 24-hour recalls and 10 different estimation methods by gender, adults aged 19 years or more, CSFII*, 1994–1996**

| Methods | Energy (kcal) | Fat (g) | Carbohydrate (g) | Fiber (g) | Vitamin A (RE*) | Vitamin C (mg) | Iron (mg) |
|---|---|---|---|---|---|---|---|
| *Men* | | | | | | | |
| Mean | | | | | | | |
| By gender and portion size | 234 | 10.8 | 31.1 | 2.22 | 248 | 23.7 | 2.89 |
| By age, gender, and portion size | 233 | 10.8 | 31.0 | 2.23 | 247 | 23.7 | 2.89 |
| By age, gender, and portion size/collapsing† | 233 | 10.8 | 30.9 | 2.22 | 248 | 23.6 | 2.89 |
| Regression‡ | 232 | 10.8 | 30.9 | 2.22 | 251 | 23.7 | 2.89 |
| Regression excluding outliers§ | 223 | 10.6 | 29.2 | 2.19 | 237 | 22.8 | 2.80 |
| Median | | | | | | | |
| By gender and portion size | 242 | 11.2 | 31.0 | 2.32 | 242 | 23.5 | 2.93 |
| By age, gender, and portion size | 239 | 11.2 | 30.7 | 2.33 | 244 | 23.3 | 2.92 |
| By age, gender, and portion size/collapsing† | 239 | 11.1 | 30.8 | 2.32 | 243 | 23.3 | 2.91 |
| Regression‡ | 239 | 11.1 | 30.6 | 2.31 | 249 | 23.4 | 2.92 |
| Block¶ | 244 | 11.6 | 32.1 | 2.38 | 254 | 24.1 | 3.03 |
| *Women* | | | | | | | |
| Mean | | | | | | | |
| By gender and portion size | 125 | 6.5 | 18.2 | 1.55 | 193 | 18.0 | 1.93 |
| By age, gender, and portion size | 125 | 6.5 | 18.1 | 1.58 | 200 | 18.1 | 1.92 |
| By age, gender, and portion size/collapsing† | 124 | 6.5 | 18.0 | 1.56 | 195 | 18.0 | 1.92 |
| Regression‡ | 124 | 6.5 | 18.1 | 1.56 | 196 | 18.1 | 1.92 |
| Regression excluding outliers§ | 123 | 6.4 | 17.4 | 1.55 | 187 | 17.4 | 1.89 |
| Median | | | | | | | |
| By gender and portion size | 131 | 6.7 | 18.1 | 1.63 | 189 | 18.2 | 1.99 |
| By age, gender, and portion size | 129 | 6.7 | 18.0 | 1.65 | 196 | 18.3 | 1.96 |
| By age, gender, and portion size/collapsing† | 129 | 6.7 | 18.0 | 1.65 | 191 | 18.3 | 1.95 |
| Regression‡ | 129 | 6.7 | 18.0 | 1.64 | 190 | 18.2 | 1.96 |
| Block¶ | 132 | 6.9 | 18.8 | 1.70 | 198 | 19.0 | 2.04 |

* CSFII, Continuing Survey of Food Intakes by Individuals; RE, retinol equivalents.

† If there were fewer than 10 individuals within any age, gender, and portion size strata, adjacent age group strata were collapsed.

‡ Regression model: nutrient intake = $\beta_0 + \beta_1$ Age2 + $\beta_2$ Age3 + $\beta_3$ Size2 + $\beta_4$ Size3 + $\varepsilon$, where $\text{Age}_i$ is an indicator for age group *i* and $\text{Size}_i$ is an indicator for portion-size group *i*.

§ Outliers are defined as observations having squared errors greater than three times the mean squared error.

¶ Weighted median nutrient density × age-gender-portion size-specific median gram weights.

adult men and women, age group has no appreciable impact on nutrient estimation. This suggests that investigators could simplify approaches to FFQ nutrient database development by excluding age as a factor. Further, grouping individuals who consume a food or foods into gender- and portion size-specific versus age-, gender-, and portion size-specific categories leads to cell sizes that are more likely to provide stable nutrient estimates.

An obvious question is, "Which is the optimal data-driven method for creating a nutrient database for an FFQ?" These data suggest that mean methods are best, but among the mean methods, none is clearly superior. The mean regression method excluding outliers was best overall in terms of mean absolute error, but performed less well with respect to mean error (bias) and mean squared error. Therefore, it is difficult to pick any single mean method over another because it is unclear whether it is better to have many estimates off by a little or a few estimates off by a lot when relating nutrients to disease outcomes. However, differences between any of the mean methods were small at best, and all performed quite well. This being the case, it makes sense to consider which method is the simplest and easiest to use, and that, we conclude, is the portion size × gender method.

Many FFQs do not query portion size. The data from this research suggest that in developing a nutrient database for such an FFQ, a mean rather than a median method should be used (excluding portion size). Further research is necessary,

**TABLE 4. Mean squared error between reported nutrient intakes from 24-hour recalls and 10 different estimation methods by gender, adults aged 19 years or more, CSFII*, 1994–1996**

| Methods | Energy ($kcal^2$) | Fat ($g^2$) | Carbo-hydrate ($g^2$) | Fiber ($g^2$) | Vitamin A ($RE^2$*) | Vitamin C ($mg^2$) | Iron ($mg^2$) |
|---|---|---|---|---|---|---|---|
| *Men* | | | | | | | |
| Mean | | | | | | | |
| By gender and portion size | 133,044 | 274 | 2,422 | 11.1 | 242,259 | 2,118 | 25.5 |
| By age, gender, and portion size | 131,724 | 272 | 2,417 | 11.2 | 268,542 | 2,117 | 24.6 |
| By age, gender, and portion size/collapsing† | 131,675 | 270 | 2,417 | 11.1 | 251,563 | 2,112 | 24.8 |
| Regression‡ | 131,831 | 271 | 2,411 | 11.0 | 277,294 | 2,118 | 24.8 |
| Regression excluding outliers§ | 144,093 | 300 | 2,540 | 11.8 | 240,474 | 2,198 | 25.5 |
| Median | | | | | | | |
| By gender and portion size | 175,589 | 346 | 2,980 | 13.5 | 262,447 | 2,576 | 29.4 |
| By age, gender, and portion size | 168,694 | 342 | 2,902 | 13.4 | 288,645 | 2,483 | 27.9 |
| By age, gender, and portion size/collapsing† | 168,817 | 339 | 2,913 | 13.3 | 264,382 | 2,495 | 27.9 |
| Regression‡ | 171,589 | 343 | 2,918 | 13.4 | 328,648 | 2,529 | 28.9 |
| Block¶ | 172,259 | 365 | 3,096 | 13.6 | 377,095 | 2,598 | 29.9 |
| *Women* | | | | | | | |
| Mean | | | | | | | |
| By gender and portion size | 35,408 | 95 | 817 | 5.4 | 193,791 | 1,115 | 10.8 |
| By age, gender, and portion size | 34,964 | 95 | 811 | 5.5 | 221,454 | 1,114 | 10.6 |
| By age, gender, and portion size/collapsing† | 34,765 | 95 | 807 | 5.4 | 195,974 | 1,109 | 10.6 |
| Regression‡ | 34,840 | 95 | 807 | 5.5 | 198,527 | 1,108 | 10.6 |
| Regression excluding outliers§ | 37,604 | 101 | 834 | 5.7 | 190,429 | 1,144 | 10.8 |
| Median | | | | | | | |
| By gender and portion size | 43,322 | 115 | 939 | 6.4 | 200,691 | 1,357 | 12.6 |
| By age, gender, and portion size | 41,814 | 114 | 906 | 6.5 | 209,165 | 1,352 | 12.1 |
| By age, gender, and portion size/collapsing† | 41,714 | 114 | 905 | 6.5 | 202,824 | 1,350 | 12.0 |
| Regression‡ | 42,656 | 113 | 927 | 6.6 | 202,430 | 1,342 | 12.1 |
| Block¶ | 43,202 | 116 | 979 | 6.8 | 244,373 | 1,472 | 12.7 |

* CSFII, Continuing Survey of Food Intakes by Individuals; RE, retinol equivalents.

† If there were fewer than 10 individuals within any age, gender, and portion size strata, adjacent age group strata were collapsed.

‡ Regression model: nutrient intake = $\beta_0 + \beta_1$ Age2 + $\beta_2$ Age3 + $\beta_3$ Size2 + $\beta_4$ Size3 + $\varepsilon$, where $Age_i$ is an indicator for age group *i* and $Size_i$ is an indicator for portion-size group *i*.

§ Outliers are defined as observations having squared errors greater than three times the mean squared error.

¶ Weighted median nutrient density × age-gender-portion size-specific median gram weights.

however, to clarify whether or not age is a more important factor when portion size is not considered, since age may be a proxy for portion size.

FFQs, like other dietary assessment instruments, continue to be based on self-report. Investigators using these instruments are well aware of the errors associated with them, such as under- and overreporting, misreporting, missing data, and so forth. The question is whether FFQ nutrient estimates can be improved even with the inevitable measurement error in reporting. Investigators frequently try to improve FFQs through changes in wording, formatting, ordering, and other cognitive aspects (15, 18–21). This research provides data to show that the methods used to create a nutrient database for an FFQ may offer another means of improving FFQs. The reduction in measurement error accompanying each such improvement will result in an improvement in our ability to measure diet and disease associations.

## REFERENCES

1. Block G, Hartman AM, Dresser CM, et al. A data-based approach to diet questionnaire design and testing. Am J Epidemiol 1986;124:453–69.
2. Willett WC, Sampson L, Stampfer MJ, et al. Reproducibility and validity of a semiquantitative food frequency question-

naire. Am J Epidemiol 1985;122:51–65.
3. Willett WC. Nutritional epidemiology. Oxford, England: Oxford University Press, 1998.
4. Ursin G, Ziegler RG, Subar AF, et al. Dietary patterns associated with a low-fat diet in the National Health Examination Follow-up Study: identification of potential confounders for epidemiologic analyses. Am J Epidemiol 1993;137:916–27.
5. Kristal AR, Shattuck AL, Williams AE. Food frequency questionnaires for diet intervention research. Proceedings of the 17th National Nutrient Databank conference, Baltimore, MD, 1992. Washington, DC: International Life Sciences Institute, 1994:110–25.
6. Nebling LC, Forman MR, Graubard BI, et al. The impact of lifestyle characteristics on carotenoid intake in the United States: the 1987 National Health Interview Survey. Am J Public Health 1997;87:268–71.
7. National Research Council. Diet and health. Washington, DC: National Academy Press, 1989.
8. Buzzard IM, Sievert YA. Research priorities and recommendations for dietary assessment methodology. Am J Clin Nutr 1994;59 (Suppl.):275S–80S.
9. Tippett KS, Cypel YS, eds. Design and operation: the continuing survey of food intakes by individuals and the diet and health knowledge survey, 1994–96. Continuing survey of food intakes by individuals 1994–96, nationwide food surveys. Report 96–1. US Department of Agriculture, Agricultural Research Service, 1997.
10. Block G, Dresser CM, Hartman AM, et al. Nutrient sources in the American diet: quantitative data from the NHANES II survey. I. Vitamins and minerals. Am J Epidemiol 1985;122:13–26.
11. Block G, Dresser CM, Hartman AM, et al. Nutrient sources in the American diet: quantitative data from the NHANES II survey. II. Macronutrients and fats. Am J Epidemiol 1985;122:27–40.
12. Thompson FE. Sowers MF, Frongillo EA, et al. Sources of fiber and fat in diets of US women aged 19–50: implications for nutrition education and policy. Am J Public Health 1992;82:695–702.
13. Krebs-Smith SM, Cronin FJ, Haytowitz DB, et al. Food sources of energy, macronutrients, cholesterol, and fiber in diets of women. J Am Diet Assoc 1992;92:168–74.
14. Subar AF, Krebs-Smith SM, Cook A, et al. Dietary sources of nutrients among US adults: 1989–91. J Am Diet Assoc 1998;98:537–47.
15. Subar AF, Thompson FE, Smith AF, et al. Improving food frequency questionnaires: a qualitative approach using cognitive interviewing. J Am Diet Assoc 1995;95:781–8.
16. National Center for Health Statistics. National health interview survey cancer epidemiology data file: data description. Washington, DC: National Center for Health Statistics, 1994.
17. National Research Council, Food and Nutrition Board. Recommended dietary allowances. 10th ed. Washington, DC: National Academy Press, 1989.
18. Worsley A. Effects of varying recall periods on reported food intakes. Appetite 1991;16:69–82.
19. Serdula M, Byers T, Coates R, et al. Assessing consumption of high-fat foods: the effect of grouping foods into single questions. Epidemiology 1992;3:503–8.
20. Kuskowska-Wolk A, Holte S, Ohlander E-M, et al. Effects of different designs and extension of a food frequency questionnaire on response rate, completeness of data and food frequency responses. Int J Epidemiol 1992;21:1144–50.
21. Tylavsky FA, Sharp GB. Misclassification of nutrient and energy intake from use of closed-ended questions in epidemiologic research. Am J Epidemiol 1995;142:342–52.